The results below are generated from an R script.

```r
#' ---
#' title: "POLI176 Final Project (SP22)"
#' author: "Christian Bankard, Beibei Du, Hugo Lopez (ranked in alphabetical order)"
#' Professor: ' Molly Roberts'
#' output: pdf_document
#' ---


# Our research question: What effects did the Vietnam War have on American domestic issues?
# We predict the most discussed themes in the state of the Union with respect to Vietnam topics: drugs,
# We are going to use both LDA and STM for the Topic Modelling.
#Load package libraries that needed for our final project
library(tidyverse)
library(tokenizers) # tokenize the data
library(quanteda) # Use this for LDA
library(quanteda.textplots)
library(stm) # Structure Topic Modelling
library(seededlda)

#Set working directory, need to change this line of code
setwd("~/Desktop/POLI176:DSC161 Text as Data/Final Project")
#Load data for speeches
metadata <- read_csv("SOTU_WithText.csv")

## Rows:  236 Columns:  5- Column specification ---------------------------------
## Delimiter:  ","
## chr (4):  president, party, sotu_type, text
## dbl (1):  year
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.

# Since our focus is on Vietnam War, we need to limit the year range
# Find out the `year` range of Vietnam War
typeof(metadata$year)

## [1] "double"

class(metadata$year) # numeric

## [1] "numeric"

# select the correct years and store it into the war df
war <- metadata[which((metadata$year>=1955) & (metadata$year<=1975)),c(colnames(metadata))] # from 1 Nov

# Preprocessing of the  data and make it corpus
corpus_sotu <- corpus(war, text_field = "text") # make it into a text corpus
corpus_sotu # take a look at it

## Corpus consisting of 24 documents and 4 docvars.
## text1 :
## "Mr. President, Mr. Speaker, Members of the Congress:  First,..."
##
## text2 :
## " [Recorded on film and tape and broadcast the same day]  My ..."
```

1

```
## 
## text3 :
## " [Read before a joint session by a clerk of the House of Rep..."
## 
## text4 :
## "To the Congress of the United States:  I appear before the C..."
## 
## text5 :
## "Mr. President, Mr. Speaker, Members of the 85th Congress:  I..."
## 
## text6 :
## " [Delivered in person before a joint session]  Mr. President..."
## 
## [ reached max_ndoc ... 18 more documents ]

#Some common pre-processing, remove the punctuation and numbers
toks <- tokens(corpus_sotu, remove_punct = TRUE, remove_numbers=TRUE) # remove punctuations and numbers
toks <- tokens_wordstem(toks) # tokenize into the wordstem
toks <- tokens_select(toks,  stopwords("en"), selection = "remove") # remove the stopwords in English
dfm <- dfm(toks) # make it the document feature matrix
dfm

## Document-feature matrix of: 24 documents, 5,628 features (79.97% sparse) and 4 docvars.
##        features
## docs    mr presid speaker member congress first extend cordial greet 84th
##   text1  2      3       1      2       24    13      5       1     1    2
##   text2  0      0       0      0        6     5      0       0     0    0
##   text3  0      0       0      1       29     2      5       0     0    0
##   text4  0      0       0      1       18     7      1       0     0    0
##   text5  2      2       1      3       10     7      2       0     1    0
##   text6  2      3       1      3       16     2      2       0     0    0
## [ reached max_ndoc ... 18 more documents, reached max_nfeat ... 5,618 more features ]

#Create a document feature matrix (dfm) and trim it with words appeared at least 5%
toks <- corpus_sotu %>%
  tokens()
dfm <- dfm(toks)
dfm_trimmed <- dfm_trim(dfm, min_docfreq = 0.05, docfreq_type = "prop") # trim it and remove the words
dfm_trimmed # the cleaned up document matrix

## Document-feature matrix of: 24 documents, 5,099 features (72.87% sparse) and 4 docvars.
##        features
## docs    mr    . president   , speaker members  of the congress   :
##   text1  2 333         3 368       1       2 370 440      24 11
##   text2  0  52         0  44       0       0  43  79       6  2
##   text3  0 381         0 336       0       0 420 609      29 10
##   text4  0 190         0 178       0       1 207 278      18  6
##   text5  2 242         2 260       1       3 231 312      10 22
##   text6  2 260         3 198       1       2 239 291      16  8
## [ reached max_ndoc ... 18 more documents, reached max_nfeat ... 5,089 more features ]

# the word cloud for documents from the 20 years of Vietnam War
textplot_wordcloud(dfm_trimmed, col="black")
```
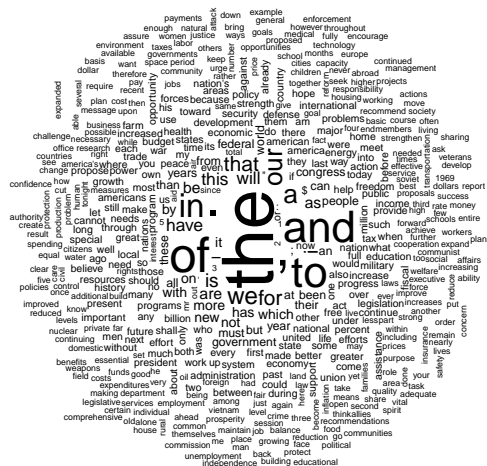
```
############# LDA #############

#Run LDA using quanteda
lda <- textmodel_lda(dfm_trimmed, k = 10)
#Most likely term for each topic
lda.terms <- terms(lda, 10) # check the top words for each topic
lda.terms

##       topic1       topic2    topic3      topic4        topic5      topic6     topic7
## [1,] "must"       "-"       "-"         "energy"      "free"      "-"        "the"
## [2,] "government" "our"     "$"         "federal"     "world"     "also"     "."
## [3,] "justice"    "or"      "billion"   "$"           "defense"   "which"    ","
## [4,] "citizens"   "and"     "percent"   "would"       "nations"   "on"       "of"
## [5,] "recommend"  "but"     "million"   "will"        "peace"     "new"      "and"
## [6,] "shall"      "must"    "last"      "-"           "security"  "for"      "to"
## [7,] "human"      "not"     "very"      "oil"         "military"  "have"     "in"
## [8,] "laws"       "its"     "act"       "legislation" "all"       "reform"   "a"
## [9,] ";"          "his"     "who"       "economic"    "economic"  "my"       "our"
## [10,] "peoples"   "every"   "home"      "needs"       "freedom"   "percent"  "we"
##       topic8          topic9        topic10
## [1,]  "been"          "one"         "we"
## [2,]  "have"          "position"    "that"
## [3,]  "has"           "scientific"  "i"
## [4,]  "government"    "missiles"    "will"
## [5,]  "federal"       "if"          "have"
## [6,]  "administration" "development" "america"
## [7,]  "program"       "education"   "great"
## [8,]  "also"          "greater"     "government"
## [9,]  "system"        "missile"     "you"
## [10,] "development"   "technology"  "us"
```

```r
#Topical content matrix
mu <- lda$phi
dim(mu) #10 topics, 5099 words

## [1]    10 5099

mu[1:10,1:20]

##                   mr            .    president             ,      speaker      members
## topic1  2.480836e-05 2.480836e-05 2.480836e-05 1.265226e-03 2.480836e-05 2.480836e-05
## topic2  7.813817e-04 1.532121e-05 1.532121e-05 1.685333e-04 1.532121e-05 3.539199e-03
## topic3  2.019835e-05 2.019835e-05 2.019835e-05 2.221818e-04 2.019835e-05 2.019835e-05
## topic4  2.210211e-04 2.009283e-05 2.009283e-05 2.009283e-05 2.009283e-05 2.009283e-05
## topic5  1.735539e-05 1.735539e-05 1.735539e-05 1.735539e-05 1.909092e-04 1.735539e-05
## topic6  1.346900e-04 1.346900e-04 1.224455e-05 1.224455e-05 1.224455e-05 1.224455e-05
## topic7  4.653034e-05 6.190451e-02 2.198787e-04 6.032613e-02 9.123596e-07 4.479686e-04
## topic8  1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05
## topic9  1.330348e-03 3.244752e-05 3.244752e-05 3.569227e-04 3.244752e-05 3.244752e-05
## topic10 2.678596e-03 9.204798e-06 7.280995e-03 1.012528e-04 2.494500e-03 9.204798e-06
##                   of          the     congress            :        first            i
## topic1  2.728919e-04 2.480836e-05 2.480836e-05 2.480836e-05 2.480836e-05 2.728919e-04
## topic2  1.532121e-05 1.532121e-05 1.532121e-05 1.685333e-04 1.532121e-05 1.394230e-03
## topic3  2.019835e-05 2.019835e-05 2.019835e-05 2.019835e-05 2.019835e-05 2.019835e-05
## topic4  2.009283e-05 2.009283e-05 2.009283e-05 3.636802e-03 2.009283e-05 2.009283e-05
## topic5  1.752894e-03 1.735539e-05 1.735539e-05 1.735539e-05 1.735539e-05 1.735539e-05
## topic6  1.224455e-05 1.224455e-05 1.224455e-05 1.224455e-05 1.224455e-05 3.795810e-04
## topic7  5.537202e-02 8.568061e-02 4.006171e-03 2.208823e-03 1.843879e-03 1.002774e-02
## topic8  1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05 3.854754e-03
## topic9  3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05
## topic10 9.204798e-06 9.204798e-06 7.465091e-03 9.204798e-06 1.012528e-04 2.421782e-02
##               extend     greetings           to           we        shall         have
## topic1  1.265226e-03 2.480836e-05 2.480836e-05 2.480836e-05 7.963482e-03 2.480836e-05
## topic2  1.532121e-05 1.685333e-04 1.532121e-05 1.532121e-05 1.532121e-05 6.281696e-04
## topic3  2.019835e-05 2.019835e-05 4.241653e-04 2.019835e-05 4.241653e-04 2.019835e-05
## topic4  4.219494e-04 2.009283e-05 2.009283e-05 2.009283e-05 2.009283e-05 6.228777e-04
## topic5  1.735539e-05 1.735539e-05 5.380170e-04 1.735539e-05 1.735539e-05 1.735539e-05
## topic6  1.224455e-05 1.224455e-05 9.918084e-04 1.224455e-05 1.224455e-05 1.507304e-02
## topic7  1.468899e-04 9.123596e-07 4.627579e-02 1.725363e-02 1.469811e-03 6.825363e-03
## topic8  1.011746e-05 1.112921e-04 1.011746e-05 3.136414e-04 1.011746e-05 1.903095e-02
## topic9  3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05
## topic10 9.204798e-06 9.204798e-06 9.204798e-06 3.682839e-02 9.204798e-06 1.510507e-02
##                 much           do
## topic1  2.728919e-04 2.480836e-05
## topic2  4.749575e-04 6.281696e-04
## topic3  2.019835e-05 2.019835e-05
## topic4  2.009283e-05 2.009283e-05
## topic5  1.735539e-05 1.735539e-05
## topic6  2.583600e-03 1.224455e-05
## topic7  7.855417e-04 1.488059e-03
## topic8  1.011746e-05 1.011746e-05
## topic9  3.569227e-04 1.654823e-03
## topic10 9.204798e-06 3.691124e-03

#Most representative words in Topic 1-10 using a for loop
for (i in 1:10){
```

```
  print(mu[i,][order(mu[i,], decreasing=T)][1:10])
  print("######################################################")
}
```

```
##       must  government     justice    citizens   recommend        shall       human
## 0.012428986 0.010692401 0.009700067 0.009203900 0.008707733 0.007963482 0.007963482
##       laws           ;     peoples
## 0.006971148 0.006723064 0.006723064
## [1] "######################################################"
##          -         our          or         and         but        must         not
## 0.042148646 0.024529256 0.019473257 0.014570470 0.011506228 0.011199804 0.010586956
##        its         his       every
## 0.008748410 0.007216290 0.006603441
## [1] "######################################################"
##          -           $     billion     percent     million        last        very
## 0.053949787 0.036983175 0.019410612 0.018602678 0.012341190 0.010523339 0.009715405
##        act         who        home
## 0.009715405 0.008301521 0.007695571
## [1] "######################################################"
##     energy     federal           $       would        will           -         oil
## 0.022724989 0.014687858 0.011473005 0.010267436 0.009865579 0.009061866 0.009061866
## legislation    economic       needs
## 0.008459081 0.008258153 0.008258153
## [1] "######################################################"
##       free       world     defense     nations       peace    security    military         all
## 0.02466200 0.01615786 0.01615786 0.01563720 0.01511654 0.01442233 0.01372811 0.01251323
##   economic     freedom
## 0.01233968 0.01199257
## [1] "######################################################"
##          -        also       which          on         new         for        have
## 0.027440032 0.018623958 0.018256621 0.018134176 0.017399503 0.017277057 0.015073039
##     reform          my     percent
## 0.008093646 0.007726310 0.007726310
## [1] "######################################################"
##        the           .           ,          of         and          to          in           a
## 0.08568061 0.06190451 0.06032613 0.05537202 0.04642177 0.04627579 0.03384946 0.02401422
##        our          we
## 0.02080271 0.01725363
## [1] "######################################################"
##           been        have         has   government     federal administration
##     0.026821396   0.019030949  0.018120378    0.016400409     0.015489837     0.011746375
##        program         also      system  development
##     0.011139328   0.010835804  0.010633454    0.009621708
## [1] "######################################################"
##        one    position  scientific    missiles          if development   education
## 0.015607255 0.007495376 0.007170901 0.006846426 0.006521951 0.006197476 0.006197476
##    greater     missile  technology
## 0.005873000 0.005873000 0.005548525
## [1] "######################################################"
##         we        that           i        will        have     america       great  government
## 0.03682839 0.03544768 0.02421782 0.01611760 0.01510507 0.01308002 0.01298797 0.01289592
##        you          us
## 0.01280387 0.01188339
## [1] "######################################################"
```

```r
#Topical prevalence matrix
pi <- lda$theta
dim(pi) #24 10
```

```
## [1] 24 10
```

```r
#Most representative documents in Topic 1-10
for (i in 1:10){
  print(war[order(pi[i,],decreasing=T),])
  print("###################################################")
}
```

```
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  3 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  4 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  6 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
##  7 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
##  8 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  9 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
## 10 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
## [1] "###################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  3 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  4 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  5 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  6 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  7 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
##  8 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  9 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
## 10 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
## [1] "###################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  3 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  4 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  6 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
##  7 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  8 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
##  9 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
## 10 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
## [1] "###################################################"
## # A tibble: 10 x 5
```

```
##    president             year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  3 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  4 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  5 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  6 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
##  7 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  8 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
##  9 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
## 10 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  3 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  4 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  5 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  6 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  7 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  8 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
##  9 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
## 10 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  3 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  4 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  6 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  7 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  8 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
##  9 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
## 10 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  3 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  4 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  6 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  7 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  8 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
##  9 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
## 10 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
```

```
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  3 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  4 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  6 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  7 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
##  8 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
##  9 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
## 10 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  3 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  4 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  5 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
##  6 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  7 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
##  8 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  9 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
## 10 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  3 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
##  4 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  6 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  7 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  8 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
##  9 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
## 10 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
## [1] "#####################################################"

############ STM ###########
# Get the `clusterFightinWords` function from Discussion3.R; We will only define this function once
clusterFightinWords <- function(dfm, clust.vect, alpha.0=100) {
  overall.terms <- colSums(dfm)
  # n and n_k in Monroe et al.
  n <- sum(overall.terms)
  # alpha_{kw} in Monroe et al.
  prior.terms <- overall.terms / n * alpha.0
  # y_{kw}(i) in Monroe et al.
  cluster.terms <- colSums(dfm[clust.vect, ])
```

```r
  # n_k(i) in Monroe et al.
  cluster.n <- sum(cluster.terms)
  cluster.term.odds <-
    (cluster.terms + prior.terms) /
    (cluster.n + alpha.0 - cluster.terms - prior.terms)
  overall.term.odds <-
    (overall.terms + prior.terms) /
    (n + alpha.0 - overall.terms - prior.terms)
  log.odds <- log(cluster.term.odds) - log(overall.term.odds)
  variance <- 1/(cluster.terms + prior.terms) + 1/(overall.terms + prior.terms)
  output <- log.odds / sqrt(variance)
  names(output) <- colnames(dfm)
  return(output)
}
# Find words that are distinctive between newsletters written by Democrats and Republicans
terms <- clusterFightinWords(dfm_trimmed,war$sotu_type == "speech")
sort(terms, decreasing=T)[1:10] #nation,people

##       we       you        is  america              ,       if      let   nation   people  nations
## 4.228220 3.483241 3.336494 3.147572 3.098360 2.808781 2.671044 2.655983 2.594946 2.578803

terms <- clusterFightinWords(dfm_trimmed,war$sotu_type == "written")
sort(terms, decreasing=T)[1:10] #federal, administration, veterans, education

##           been          also       federal administration              has          have
##       6.690192      5.119512      4.309598      4.257529         3.983291      3.781330
##       veterans        during          1973     education
##       3.529462      3.420720      3.240436      3.231743

# Structural Topic Model
#STM
#library(tm)
temp<-textProcessor(documents=war$text,metadata=war)

## Building corpus...
## Converting to Lower Case...
## Removing punctuation...
## Removing stopwords...
## Removing numbers...
## Stemming...
## Creating Output...

out <- prepDocuments(temp$documents, temp$vocab, temp$meta)

## Removing 3101 of 6301 terms (3101 of 27073 tokens) due to frequency
## Your corpus now has 24 documents, 3200 terms and 23972 tokens.

# Run the model
# Should we change the value of K
model.stm <- stm(out$documents, out$vocab, K = 10, prevalence = ~ party,
                 data = out$meta, max.em.its = 10)

## Beginning Spectral Initialization
##    Calculating the gram matrix...
##    Finding anchor words...
##    ..........
```

9

```
##   Recovering initialization...
##   ..............................
## Initialization complete.
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 1 (approx. per word bound = -6.843)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 2 (approx. per word bound = -6.769, relative change = 1.077e-02)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 3 (approx. per word bound = -6.740, relative change = 4.206e-03)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 4 (approx. per word bound = -6.727, relative change = 1.915e-03)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 5 (approx. per word bound = -6.720, relative change = 1.052e-03)
## Topic 1: will, program, year, new, nation
##  Topic 2: will, new, govern, can, peopl
##  Topic 3: must, nation, will, year, can
##  Topic 4: will, year, feder, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, world, will, must, peac
##  Topic 7: will, nation, new, year, world
##  Topic 8: year, will, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, will, year, program, must
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 6 (approx. per word bound = -6.716, relative change = 6.577e-04)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 7 (approx. per word bound = -6.713, relative change = 4.550e-04)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 8 (approx. per word bound = -6.711, relative change = 3.559e-04)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 9 (approx. per word bound = -6.708, relative change = 3.055e-04)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Model Terminated Before Convergence Reached
```

```
model.stm <- stm(out$documents, out$vocab, K = 10, prevalence = ~ party,
                 data = out$meta)

## Beginning Spectral Initialization
##    Calculating the gram matrix...
##    Finding anchor words...
##     ..........
##    Recovering initialization...
##     ..............................
## Initialization complete.
##     .......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 1 (approx. per word bound = -6.843)
##     .......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 2 (approx. per word bound = -6.769, relative change = 1.077e-02)
##     .......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 3 (approx. per word bound = -6.740, relative change = 4.206e-03)
##     .......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 4 (approx. per word bound = -6.727, relative change = 1.915e-03)
##     .......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 5 (approx. per word bound = -6.720, relative change = 1.052e-03)
## Topic 1: will, program, year, new, nation
##   Topic 2: will, new, govern, can, peopl
##   Topic 3: must, nation, will, year, can
##   Topic 4: will, year, feder, new, program
##   Topic 5: year, will, congress, nation, can
##   Topic 6: nation, world, will, must, peac
##   Topic 7: will, nation, new, year, world
##   Topic 8: year, will, america, nation, american
##   Topic 9: will, nation, year, peopl, must
##   Topic 10: nation, will, year, program, must
##     .......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 6 (approx. per word bound = -6.716, relative change = 6.577e-04)
##     .......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 7 (approx. per word bound = -6.713, relative change = 4.550e-04)
##     .......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 8 (approx. per word bound = -6.711, relative change = 3.559e-04)
##     .......................
## Completed E-Step (0 seconds).
```

```
## Completed M-Step.
## Completing Iteration 9 (approx. per word bound = -6.708, relative change = 3.055e-04)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 10 (approx. per word bound = -6.707, relative change = 2.648e-04)
## Topic 1: will, program, year, new, nation
##  Topic 2: will, new, can, govern, peopl
##  Topic 3: nation, must, will, year, can
##  Topic 4: will, year, feder, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, world, will, must, year
##  Topic 7: will, nation, new, year, world
##  Topic 8: will, year, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, will, program, year, govern
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 11 (approx. per word bound = -6.705, relative change = 2.049e-04)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 12 (approx. per word bound = -6.704, relative change = 1.551e-04)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 13 (approx. per word bound = -6.704, relative change = 1.074e-04)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 14 (approx. per word bound = -6.703, relative change = 9.148e-05)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 15 (approx. per word bound = -6.702, relative change = 9.105e-05)
## Topic 1: will, program, year, new, last
##  Topic 2: will, new, can, govern, peopl
##  Topic 3: nation, will, must, year, world
##  Topic 4: will, year, feder, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, world, will, must, year
##  Topic 7: will, nation, new, year, world
##  Topic 8: will, year, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, will, program, year, govern
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 16 (approx. per word bound = -6.702, relative change = 6.806e-05)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 17 (approx. per word bound = -6.701, relative change = 5.998e-05)
```

```
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 18 (approx. per word bound = -6.701, relative change = 6.022e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 19 (approx. per word bound = -6.701, relative change = 5.880e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 20 (approx. per word bound = -6.700, relative change = 5.545e-05)
## Topic 1: will, program, year, new, also
##  Topic 2: will, new, can, govern, peopl
##  Topic 3: nation, will, must, year, world
##  Topic 4: will, year, feder, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, world, will, must, year
##  Topic 7: will, nation, new, year, world
##  Topic 8: will, year, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, will, program, year, govern
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 21 (approx. per word bound = -6.700, relative change = 5.017e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 22 (approx. per word bound = -6.700, relative change = 4.529e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 23 (approx. per word bound = -6.699, relative change = 4.172e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 24 (approx. per word bound = -6.699, relative change = 3.892e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 25 (approx. per word bound = -6.699, relative change = 3.270e-05)
## Topic 1: will, program, year, new, also
##  Topic 2: will, new, govern, peopl, can
##  Topic 3: nation, will, must, year, world
##  Topic 4: will, year, feder, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, will, world, must, year
##  Topic 7: will, nation, new, year, world
##  Topic 8: will, year, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, program, will, year, govern
## ........................
## Completed E-Step (0 seconds).
```

```
## Completed M-Step.
## Completing Iteration 26 (approx. per word bound = -6.699, relative change = 3.062e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 27 (approx. per word bound = -6.698, relative change = 3.205e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 28 (approx. per word bound = -6.698, relative change = 3.032e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 29 (approx. per word bound = -6.698, relative change = 2.073e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 30 (approx. per word bound = -6.698, relative change = 1.889e-05)
## Topic 1: will, program, year, new, also
##  Topic 2: will, new, govern, peopl, can
##  Topic 3: nation, will, must, year, world
##  Topic 4: will, year, feder, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, will, world, must, year
##  Topic 7: will, nation, new, year, world
##  Topic 8: will, year, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, program, will, year, govern
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 31 (approx. per word bound = -6.698, relative change = 1.864e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 32 (approx. per word bound = -6.698, relative change = 1.940e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 33 (approx. per word bound = -6.698, relative change = 2.044e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 34 (approx. per word bound = -6.697, relative change = 2.077e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 35 (approx. per word bound = -6.697, relative change = 2.147e-05)
## Topic 1: will, program, year, new, also
##  Topic 2: will, govern, new, peopl, can
##  Topic 3: nation, will, must, year, world
##  Topic 4: will, feder, year, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, will, world, must, year
```

```
##  Topic 7: will, nation, new, year, world
##  Topic 8: will, year, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, program, will, year, govern
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 36 (approx. per word bound = -6.697, relative change = 2.156e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 37 (approx. per word bound = -6.697, relative change = 2.090e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 38 (approx. per word bound = -6.697, relative change = 2.022e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 39 (approx. per word bound = -6.697, relative change = 1.892e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 40 (approx. per word bound = -6.697, relative change = 1.891e-05)
## Topic 1: will, program, year, new, also
##  Topic 2: will, govern, new, peopl, can
##  Topic 3: nation, will, must, year, world
##  Topic 4: will, feder, year, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, will, world, must, year
##  Topic 7: will, nation, new, year, world
##  Topic 8: will, year, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, program, will, year, govern
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 41 (approx. per word bound = -6.697, relative change = 1.790e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 42 (approx. per word bound = -6.696, relative change = 1.729e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 43 (approx. per word bound = -6.696, relative change = 1.598e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 44 (approx. per word bound = -6.696, relative change = 1.561e-05)
## ........................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 45 (approx. per word bound = -6.696, relative change = 1.539e-05)
```

```
## Topic 1: will, year, program, new, also
##  Topic 2: will, govern, peopl, new, can
##  Topic 3: nation, will, must, year, world
##  Topic 4: will, feder, year, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, will, world, must, year
##  Topic 7: will, nation, new, year, world
##  Topic 8: will, year, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, program, will, year, govern
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 46 (approx. per word bound = -6.696, relative change = 1.469e-05)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 47 (approx. per word bound = -6.696, relative change = 1.347e-05)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 48 (approx. per word bound = -6.696, relative change = 1.228e-05)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 49 (approx. per word bound = -6.696, relative change = 1.240e-05)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 50 (approx. per word bound = -6.696, relative change = 1.192e-05)
## Topic 1: will, year, program, new, also
##  Topic 2: will, govern, peopl, new, can
##  Topic 3: nation, will, must, year, world
##  Topic 4: will, feder, year, new, program
##  Topic 5: year, will, congress, nation, can
##  Topic 6: nation, will, world, must, year
##  Topic 7: will, nation, new, year, world
##  Topic 8: will, year, america, nation, american
##  Topic 9: will, nation, year, peopl, must
##  Topic 10: nation, program, will, year, govern
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 51 (approx. per word bound = -6.696, relative change = 1.183e-05)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 52 (approx. per word bound = -6.696, relative change = 1.139e-05)
## ......................
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 53 (approx. per word bound = -6.695, relative change = 1.039e-05)
## ......................
## Completed E-Step (0 seconds).
```
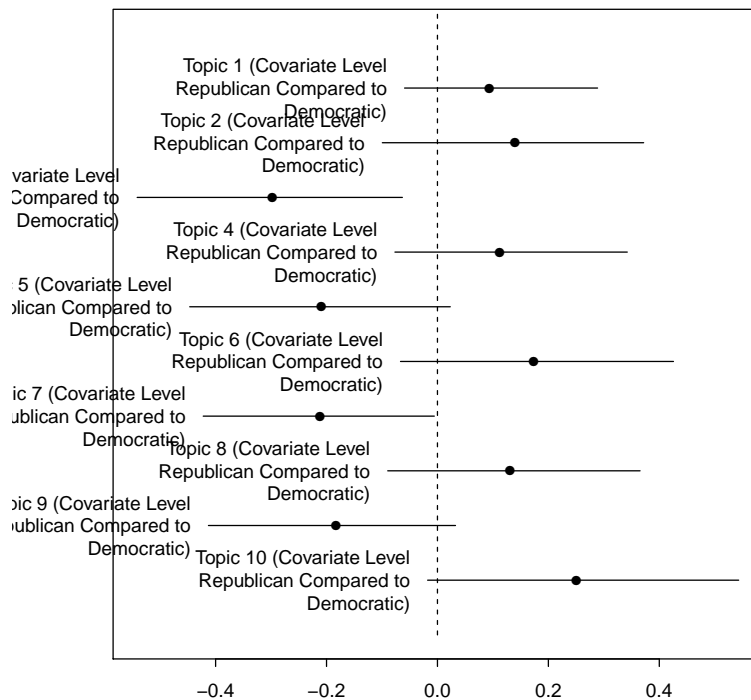
```
## Completed M-Step.
## Model Converged

labelTopics(model.stm)

## Topic 1 Top Words:
##     Highest Prob: will, year, program, new, also, can, last
##     FREX: substanti, expand, also, older, influenc, research, career
##     Lift: modest, network, alleg, diversifi, film, index, manifold
##     Score: film, career, vietnam, accid, agenda, mention, hire
## Topic 2 Top Words:
##     Highest Prob: will, govern, peopl, new, america, can, nation
##     FREX: chanc, open, reform, clean, govern, let, america
##     Lift: disagre, patron, burst, nation-, peoplei, runaway, philadelphia
##     Score: seventi, disagre, genius, tonight, clean, vision, weak
## Topic 3 Top Words:
##     Highest Prob: nation, will, must, year, world, can, new
##     FREX: allianc, lack, short, nuclear, reduct, caus, cuba
##     Lift: -hour, boom, buri, camp, diffus, full-tim, lull
##     Score: full-tim, cold, cuba, viet-nam, anew, gold, west
## Topic 4 Top Words:
##     Highest Prob: will, feder, year, new, program, increas, nation
##     FREX: environment, oil, energi, feder, fund, suppli, reform
##     Lift: petroleum, automat, automot, coal, decontrol, disloc, mileag
##     Score: environment, oil, bicentenni, petroleum, veteran, transfer, privaci
## Topic 5 Top Words:
##     Highest Prob: year, will, congress, nation, can, american, billion
##     FREX: think, pass, tonight, believ, want, billion, senat
##     Lift: chairman, taxabl, told, middle-incom, chamber-, sooner, licens
##     Score: tonight, taxabl, vietnam, medicar, chairman, told, earliest
## Topic 6 Top Words:
##     Highest Prob: nation, will, world, must, year, can, peopl
##     FREX: missil, scientif, necessari, ballist, militari, expenditur, recent
##     Lift: compart, newly-develop, aggreg, annihil, aria, evidenc, foundation-ston
##     Score: ballist, compart, steel, possess, destruct, intellectu, missil
## Topic 7 Top Words:
##     Highest Prob: will, nation, new, year, world, american, can
##     FREX: man, common, atlant, europ, uniti, seek, bill
##     Lift: landscap, spontan, restless, insid, seed, elector, non-nuclear
##     Score: landscap, tonight, enlarg, restless, beauti, spontan, scholarship
## Topic 8 Top Words:
##     Highest Prob: will, year, america, nation, american, new, congress
##     FREX: look, america, energi, let, meet, presid, one
##     Lift: cardin, worldth, freshman, applaus, recal, campus, worldand
##     Score: cardin, tonight, privaci, recal, chamber, colleagu, worldth
## Topic 9 Top Words:
##     Highest Prob: will, nation, year, peopl, must, can, american
##     FREX: vietnam, tri, tonight, south, conflict, attack, north
##     Lift: -tax, outlaw, soldier, inward, perceiv, prolifer, asian
##     Score: tonight, vietnam, soldier, vietnames, asian, -tax, outlaw
## Topic 10 Top Words:
##     Highest Prob: nation, program, will, year, govern, must, feder
##     FREX: atom, project, agricultur, vigor, eight, vital, construct
##     Lift: aviat, colorado, contributori, entiti, inalien, iran, longshoremen
##     Score: immigr, entiti, veteran, statehood, survivor, atom, old-ag
```
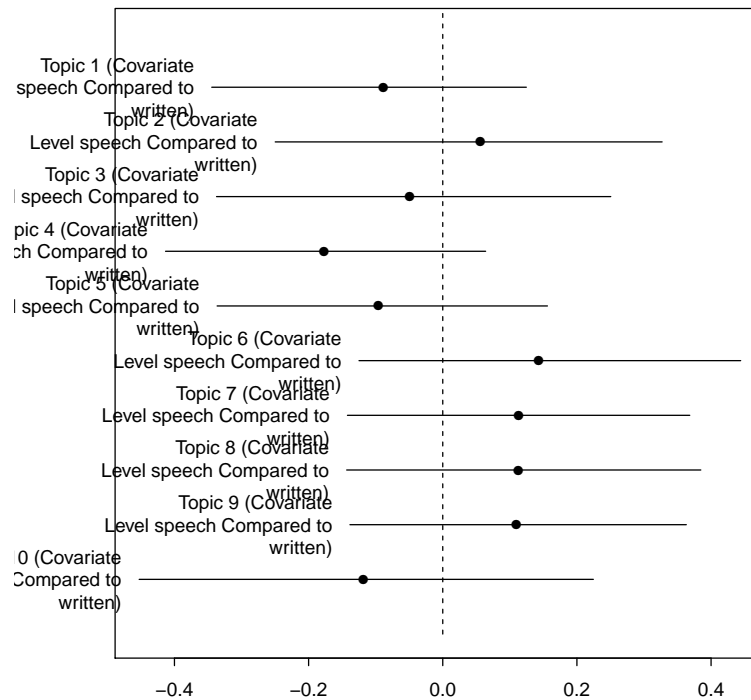
```
# topics 1-10
# difference between different parties
model.stm.ee <- estimateEffect(1:10 ~ party, model.stm, meta = out$meta)
# right associated with cov.value1
plot(model.stm.ee, "party", method="difference", cov.value1="Republican", cov.value2="Democratic")
```



```
# differences between two types
model.stm.ee <- estimateEffect(1:10 ~ sotu_type, model.stm, meta = out$meta)
plot(model.stm.ee, "sotu_type", method="difference", cov.value1="speech", cov.value2="written")
```

Topic 1 (Covariate speech Compared to written)
Topic 2 (Covariate Level speech Compared to written)
Topic 3 (Covariate speech Compared to written)
Topic 4 (Covariate ch Compared to written)
Topic 5 (Covariate speech Compared to written)
Topic 6 (Covariate Level speech Compared to written)
Topic 7 (Covariate Level speech Compared to written)
Topic 8 (Covariate Level speech Compared to written)
Topic 9 (Covariate Level speech Compared to written)
0 (Covariate Compared to written)

−0.4 −0.2 0.0 0.2 0.4

```r
# The above code is the 20 years data, next we need to find out the first ten years and last ten years
######################################### first ten year of Vietnam War ##########################
# The following code relates to the first eleven/half year of the Vietnam War

war_first_ten <- metadata[which((metadata$year>=1955) & (metadata$year<=1965)),c(colnames(metadata))]
# Preprocessing of the  data and make it corpus
corpus_sotu_first_ten <- corpus(war_first_ten, text_field = "text")
corpus_sotu_first_ten
```

```
## Corpus consisting of 13 documents and 4 docvars.
## text1 :
## "Mr. President, Mr. Speaker, Members of the Congress:  First,..."
##
## text2 :
## " [Recorded on film and tape and broadcast the same day]  My ..."
##
## text3 :
## " [Read before a joint session by a clerk of the House of Rep..."
##
## text4 :
## "To the Congress of the United States:  I appear before the C..."
##
## text5 :
## "Mr. President, Mr. Speaker, Members of the 85th Congress:  I..."
##
## text6 :
## " [Delivered in person before a joint session]  Mr. President..."
##
## [ reached max_ndoc ... 7 more documents ]
```

```r
#Some common pre-processing, remove the punctuation and numbers
```

```r
toks_first_ten <- tokens(corpus_sotu_first_ten, remove_punct = TRUE, remove_numbers=TRUE)
toks_first_ten <- tokens_wordstem(toks_first_ten)
toks_first_ten <- tokens_select(toks_first_ten,  stopwords("en"), selection = "remove")
dfm <- dfm(toks_first_ten)
dfm

## Document-feature matrix of: 13 documents, 4,155 features (73.42% sparse) and 4 docvars.
##        features
## docs    mr presid speaker member congress first extend cordial greet 84th
##    text1  2      3       1      2       24    13      5       1     1    2
##    text2  0      0       0      0        6     5      0       0     0    0
##    text3  0      0       0      1       29     2      5       0     0    0
##    text4  0      0       0      1       18     7      1       0     0    0
##    text5  2      2       1      3       10     7      2       0     1    0
##    text6  2      3       1      3       16     2      2       0     0    0
## [ reached max_ndoc ... 7 more documents, reached max_nfeat ... 4,145 more features ]

#Create a document feature matrix (dfm) and trim it with words appeared at least 5%
toks_first_ten <- corpus_sotu %>%
  tokens()
dfm_first_ten <- dfm(toks_first_ten)
dfm_trimmed_first_ten <- dfm_trim(dfm_first_ten, min_docfreq = 0.05, docfreq_type = "prop")
dfm_trimmed_first_ten

## Document-feature matrix of: 24 documents, 5,099 features (72.87% sparse) and 4 docvars.
##        features
## docs    mr    . president    , speaker members  of the congress   :
##    text1  2 333         3 368       1       2 370 440       24 11
##    text2  0  52         0  44       0       0  43  79        6  2
##    text3  0 381         0 336       0       0 420 609       29 10
##    text4  0 190         0 178       0       1 207 278       18  6
##    text5  2 242         2 260       1       3 231 312       10 22
##    text6  2 260         3 198       1       2 239 291       16  8
## [ reached max_ndoc ... 18 more documents, reached max_nfeat ... 5,089 more features ]

# word cloud for the first ten years of the Vietnam War
textplot_wordcloud(dfm_trimmed_first_ten, col="blue")
```

```r
# Find words that are distinctive between newsletters written by Democrats and Republicans
terms_first_ten_s <- clusterFightinWords(dfm_trimmed_first_ten,war_first_ten$sotu_type == "speech")
sort(terms_first_ten_s, decreasing=T)[1:10] #nation,people
```

```
##        it        or       but       not         ,        we       oil      faith    nation   poverty
## 1.877784 1.658427 1.563160 1.560399 1.504458 1.487339 1.438306 1.402208 1.398200 1.344162
```

```r
terms_first_ten_w <- clusterFightinWords(dfm_trimmed_first_ten,war_first_ten$sotu_type == "written")
sort(terms_first_ten_w, decreasing=T)[1:10]#federal, administration, veterans, education
```

```
##      been   program      1953       has      also   sharing   welfare   revenue influence
## 5.177928  4.578795  3.740694  3.085252  2.930361  2.777721  2.718907  2.714781  2.594274
##      last
## 2.518789
```

```r
############# LDA ###########

#Run LDA using quanteda
lda <- textmodel_lda(dfm_trimmed_first_ten, k = 10)
#Most likely term for each topic
lda<- terms(lda, 10)
lda.terms
```

```
##         topic1        topic2   topic3    topic4     topic5      topic6    topic7
## [1,]  "must"        "-"      "-"       "energy"   "free"      "-"       "the"
## [2,]  "government"  "our"    "$"       "federal"  "world"     "also"    "."
## [3,]  "justice"     "or"     "billion" "$"        "defense"   "which"   ","
## [4,]  "citizens"    "and"    "percent" "would"    "nations"   "on"      "of"
## [5,]  "recommend"   "but"    "million" "will"     "peace"     "new"     "and"
## [6,]  "shall"       "must"   "last"    "-"        "security"  "for"     "to"
## [7,]  "human"       "not"    "very"    "oil"      "military"  "have"    "in"
```

```
##  [8,] "laws"           "its"    "act"    "legislation" "all"        "reform"   "a"
##  [9,] ";"              "his"    "who"    "economic"    "economic"   "my"       "our"
## [10,] "peoples"        "every"  "home"   "needs"       "freedom"    "percent"  "we"
##        topic8           topic9        topic10
##  [1,] "been"           "one"         "we"
##  [2,] "have"           "position"    "that"
##  [3,] "has"            "scientific"  "i"
##  [4,] "government"     "missiles"    "will"
##  [5,] "federal"        "if"          "have"
##  [6,] "administration" "development" "america"
##  [7,] "program"        "education"   "great"
##  [8,] "also"           "greater"     "government"
##  [9,] "system"         "missile"     "you"
## [10,] "development"    "technology"  "us"
```

```r
#Topical content matrix
mu <- lda$phi
```

```
## Error in lda$phi: $ operator is invalid for atomic vectors
```

```r
dim(mu) #10 topics, 5099 words
```

```
## [1]   10 5099
```

```r
mu[1:10,1:20]
```

```
##                  mr           .    president            ,      speaker      members
## topic1  2.480836e-05 2.480836e-05 2.480836e-05 1.265226e-03 2.480836e-05 2.480836e-05
## topic2  7.813817e-04 1.532121e-05 1.532121e-05 1.685333e-04 1.532121e-05 3.539199e-03
## topic3  2.019835e-05 2.019835e-05 2.019835e-05 2.221818e-04 2.019835e-05 2.019835e-05
## topic4  2.210211e-04 2.009283e-05 2.009283e-05 2.009283e-05 2.009283e-05 2.009283e-05
## topic5  1.735539e-05 1.735539e-05 1.735539e-05 1.735539e-05 1.909092e-04 1.735539e-05
## topic6  1.346900e-04 1.346900e-04 1.224455e-05 1.224455e-05 1.224455e-05 1.224455e-05
## topic7  4.653034e-05 6.190451e-02 2.198787e-04 6.032613e-02 9.123596e-07 4.479686e-04
## topic8  1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05
## topic9  1.330348e-03 3.244752e-05 3.244752e-05 3.569227e-04 3.244752e-05 3.244752e-05
## topic10 2.678596e-03 9.204798e-06 7.280995e-03 1.012528e-04 2.494500e-03 9.204798e-06
##                  of          the     congress            :        first            i
## topic1  2.728919e-04 2.480836e-05 2.480836e-05 2.480836e-05 2.480836e-05 2.728919e-04
## topic2  1.532121e-05 1.532121e-05 1.532121e-05 1.685333e-04 1.532121e-05 1.394230e-03
## topic3  2.019835e-05 2.019835e-05 2.019835e-05 2.019835e-05 2.019835e-05 2.019835e-05
## topic4  2.009283e-05 2.009283e-05 2.009283e-05 3.636802e-03 2.009283e-05 2.009283e-05
## topic5  1.752894e-03 1.735539e-05 1.735539e-05 1.735539e-05 1.735539e-05 1.735539e-05
## topic6  1.224455e-05 1.224455e-05 1.224455e-05 1.224455e-05 1.224455e-05 3.795810e-04
## topic7  5.537202e-02 8.568061e-02 4.006171e-03 2.208823e-03 1.843879e-03 1.002774e-02
## topic8  1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05 1.011746e-05 3.854754e-03
## topic9  3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05
## topic10 9.204798e-06 9.204798e-06 7.465091e-03 9.204798e-06 1.012528e-04 2.421782e-02
##              extend     greetings           to           we        shall         have
## topic1  1.265226e-03 2.480836e-05 2.480836e-05 2.480836e-05 7.963482e-03 2.480836e-05
## topic2  1.532121e-05 1.685333e-04 1.532121e-05 1.532121e-05 1.532121e-05 6.281696e-04
## topic3  2.019835e-05 2.019835e-05 4.241653e-04 2.019835e-05 4.241653e-04 2.019835e-05
## topic4  4.219494e-04 2.009283e-05 2.009283e-05 2.009283e-05 2.009283e-05 6.228777e-04
## topic5  1.735539e-05 1.735539e-05 5.380170e-04 1.735539e-05 1.735539e-05 1.735539e-05
## topic6  1.224455e-05 1.224455e-05 9.918084e-04 1.224455e-05 1.224455e-05 1.507304e-02
```

```
## topic7  1.468899e-04 9.123596e-07 4.627579e-02 1.725363e-02 1.469811e-03 6.825363e-03
## topic8  1.011746e-05 1.112921e-04 1.011746e-05 3.136414e-04 1.011746e-05 1.903095e-02
## topic9  3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05 3.244752e-05
## topic10 9.204798e-06 9.204798e-06 9.204798e-06 3.682839e-02 9.204798e-06 1.510507e-02
##                 much           do
## topic1  2.728919e-04 2.480836e-05
## topic2  4.749575e-04 6.281696e-04
## topic3  2.019835e-05 2.019835e-05
## topic4  2.009283e-05 2.009283e-05
## topic5  1.735539e-05 1.735539e-05
## topic6  2.583600e-03 1.224455e-05
## topic7  7.855417e-04 1.488059e-03
## topic8  1.011746e-05 1.011746e-05
## topic9  3.569227e-04 1.654823e-03
## topic10 9.204798e-06 3.691124e-03
```

```r
#Most representative words in Topic 1-10
for (i in 1:10){
  print(mu[i,][order(mu[i,], decreasing=T)][1:10])
  print("#######################################################")
}
```

```
##        must  government      justice    citizens   recommend       shall       human
## 0.012428986 0.010692401 0.009700067 0.009203900 0.008707733 0.007963482 0.007963482
##        laws           ;      peoples
## 0.006971148 0.006723064 0.006723064
## [1] "#######################################################"
##           -         our          or         and         but        must         not
## 0.042148646 0.024529256 0.019473257 0.014570470 0.011506228 0.011199804 0.010586956
##         its         his       every
## 0.008748410 0.007216290 0.006603441
## [1] "#######################################################"
##           -           $      billion     percent     million        last        very
## 0.053949787 0.036983175 0.019410612 0.018602678 0.012341190 0.010523339 0.009715405
##         act         who        home
## 0.009715405 0.008301521 0.007695571
## [1] "#######################################################"
##      energy     federal           $       would        will           -         oil
## 0.022724989 0.014687858 0.011473005 0.010267436 0.009865579 0.009061866 0.009061866
## legislation    economic       needs
## 0.008459081 0.008258153 0.008258153
## [1] "#######################################################"
##        free       world     defense     nations       peace    security    military         all
## 0.02466200 0.01615786 0.01615786 0.01563720 0.01511654 0.01442233 0.01372811 0.01251323
##    economic     freedom
## 0.01233968 0.01199257
## [1] "#######################################################"
##           -        also       which          on         new         for        have
## 0.027440032 0.018623958 0.018256621 0.018134176 0.017399503 0.017277057 0.015073039
##      reform          my     percent
## 0.008093646 0.007726310 0.007726310
## [1] "#######################################################"
##         the           .           ,          of         and          to          in           a
## 0.08568061 0.06190451 0.06032613 0.05537202 0.04642177 0.04627579 0.03384946 0.02401422
```

```
##        our         we
## 0.02080271 0.01725363
## [1] "#####################################################"
##         been          have           has    government      federal administration
##    0.026821396    0.019030949    0.018120378    0.016400409    0.015489837    0.011746375
##        program           also         system    development
##    0.011139328    0.010835804    0.010633454    0.009621708
## [1] "#####################################################"
##         one     position   scientific      missiles          if development     education
## 0.015607255 0.007495376 0.007170901 0.006846426 0.006521951 0.006197476 0.006197476
##      greater       missile   technology
## 0.005873000 0.005873000 0.005548525
## [1] "#####################################################"
##          we          that            i          will          have       america        great government
## 0.03682839 0.03544768 0.02421782 0.01611760 0.01510507 0.01308002 0.01298797 0.01289592
##          you            us
## 0.01280387 0.01188339
## [1] "#####################################################"
```

```
#Topical prevalence matrix
pi <- lda$theta
```

```
## Error in lda$theta:  $ operator is invalid for atomic vectors
```

```
dim(pi) #24 10
```

```
## [1] 24 10
```

```
#Most representative documents in Topic 1-10
for (i in 1:10){
  print(war_first_ten[order(pi[i,],decreasing=T),])
  print("#####################################################")
}
```

```
## # A tibble: 10 x 5
##    president            year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower 1960 Republican speech     "\n\n[Delivered in person before a joi~
##  2 John F. Kennedy      1961 Democratic written    "\n\n[ As delivered in person before a~
##  3 Dwight D. Eisenhower 1955 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  4 Dwight D. Eisenhower 1958 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  5 Dwight D. Eisenhower 1956 Republican speech     "\n\n[Recorded on film and tape and br~
##  6 Dwight D. Eisenhower 1956 Republican written    "\n\n[Read before a joint session by a~
##  7 Dwight D. Eisenhower 1959 Republican speech     "\n\n[Delivered in person before a joi~
##  8 Dwight D. Eisenhower 1961 Republican written    "\n\nTo the Congress of the United Sta~
##  9 John F. Kennedy      1962 Democratic speech     "\n\n[As delivered in person before a ~
## 10 Dwight D. Eisenhower 1957 Republican speech     "To the Congress of the United States:~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president            year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower 1960 Republican speech     "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower 1958 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  3 John F. Kennedy      1961 Democratic written    "\n\n[ As delivered in person before a~
##  4 John F. Kennedy      1962 Democratic speech     "\n\n[As delivered in person before a ~
```

```
##  5 Dwight D. Eisenhower  1955 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  6 Dwight D. Eisenhower  1956 Republican speech     "\n\n[Recorded on film and tape and br~
##  7 Dwight D. Eisenhower  1956 Republican written    "\n\n[Read before a joint session by a~
##  8 Dwight D. Eisenhower  1961 Republican written    "\n\nTo the Congress of the United Sta~
##  9 Dwight D. Eisenhower  1957 Republican speech     "To the Congress of the United States:~
## 10 Dwight D. Eisenhower  1959 Republican speech     "\n\n[Delivered in person before a joi~
## [1] "####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech     "\n\n[Delivered in person before a joi~
##  2 John F. Kennedy       1961 Democratic written    "\n\n[ As delivered in person before a~
##  3 Dwight D. Eisenhower  1955 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  4 Dwight D. Eisenhower  1958 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  5 Dwight D. Eisenhower  1956 Republican speech     "\n\n[Recorded on film and tape and br~
##  6 Dwight D. Eisenhower  1959 Republican speech     "\n\n[Delivered in person before a joi~
##  7 Dwight D. Eisenhower  1961 Republican written    "\n\nTo the Congress of the United Sta~
##  8 Dwight D. Eisenhower  1957 Republican speech     "To the Congress of the United States:~
##  9 John F. Kennedy       1962 Democratic speech     "\n\n[As delivered in person before a ~
## 10 Dwight D. Eisenhower  1956 Republican written    "\n\n[Read before a joint session by a~
## [1] "####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech     "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1958 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  3 Dwight D. Eisenhower  1955 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  4 John F. Kennedy       1961 Democratic written    "\n\n[ As delivered in person before a~
##  5 Dwight D. Eisenhower  1956 Republican speech     "\n\n[Recorded on film and tape and br~
##  6 Dwight D. Eisenhower  1957 Republican speech     "To the Congress of the United States:~
##  7 John F. Kennedy       1962 Democratic speech     "\n\n[As delivered in person before a ~
##  8 Dwight D. Eisenhower  1959 Republican speech     "\n\n[Delivered in person before a joi~
##  9 Dwight D. Eisenhower  1956 Republican written    "\n\n[Read before a joint session by a~
## 10 Dwight D. Eisenhower  1961 Republican written    "\n\nTo the Congress of the United Sta~
## [1] "####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech     "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1958 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  3 Dwight D. Eisenhower  1961 Republican written    "\n\nTo the Congress of the United Sta~
##  4 John F. Kennedy       1962 Democratic speech     "\n\n[As delivered in person before a ~
##  5 Dwight D. Eisenhower  1956 Republican speech     "\n\n[Recorded on film and tape and br~
##  6 Dwight D. Eisenhower  1955 Republican speech     "Mr. President, Mr. Speaker, Members o~
##  7 John F. Kennedy       1961 Democratic written    "\n\n[ As delivered in person before a~
##  8 Dwight D. Eisenhower  1959 Republican speech     "\n\n[Delivered in person before a joi~
##  9 Dwight D. Eisenhower  1957 Republican speech     "To the Congress of the United States:~
## 10 Dwight D. Eisenhower  1956 Republican written    "\n\n[Read before a joint session by a~
## [1] "####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech     "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1958 Republican speech     "Mr. President, Mr. Speaker, Members o~
```

```
##  3 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  4 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  6 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  7 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  8 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
##  9 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
## 10 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  3 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  4 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  6 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  7 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  8 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
##  9 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
## 10 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  3 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  4 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  6 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
##  7 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
##  8 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
##  9 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
## 10 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
##    <chr>                <dbl> <chr>      <chr>     <chr>
##  1 Dwight D. Eisenhower  1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 John F. Kennedy       1961 Democratic written   "\n\n[ As delivered in person before a~
##  3 Dwight D. Eisenhower  1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  4 Dwight D. Eisenhower  1961 Republican written   "\n\nTo the Congress of the United Sta~
##  5 Dwight D. Eisenhower  1956 Republican written   "\n\n[Read before a joint session by a~
##  6 Dwight D. Eisenhower  1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  7 Dwight D. Eisenhower  1959 Republican speech    "\n\n[Delivered in person before a joi~
##  8 Dwight D. Eisenhower  1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  9 Dwight D. Eisenhower  1957 Republican speech    "To the Congress of the United States:~
## 10 John F. Kennedy       1962 Democratic speech    "\n\n[As delivered in person before a ~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president             year party      sotu_type text
```

```
##    <chr>                 <dbl> <chr>        <chr>      <chr>
##  1 Dwight D. Eisenhower   1960 Republican speech    "\n\n[Delivered in person before a joi~
##  2 Dwight D. Eisenhower   1956 Republican speech    "\n\n[Recorded on film and tape and br~
##  3 Dwight D. Eisenhower   1956 Republican written   "\n\n[Read before a joint session by a~
##  4 Dwight D. Eisenhower   1958 Republican speech    "Mr. President, Mr. Speaker, Members o~
##  5 John F. Kennedy        1962 Democratic speech    "\n\n[As delivered in person before a ~
##  6 Dwight D. Eisenhower   1961 Republican written   "\n\nTo the Congress of the United Sta~
##  7 John F. Kennedy        1961 Democratic written   "\n\n[ As delivered in person before a~
##  8 Dwight D. Eisenhower   1959 Republican speech    "\n\n[Delivered in person before a joi~
##  9 Dwight D. Eisenhower   1957 Republican speech    "To the Congress of the United States:~
## 10 Dwight D. Eisenhower   1955 Republican speech    "Mr. President, Mr. Speaker, Members o~
## [1] "####################################################"


############ STM ###########

# Structural Topic Model
#STM
#library(tm)
temp_first_ten <-textProcessor(documents=war_first_ten$text,metadata=war_first_ten)

## Building corpus...
## Converting to Lower Case...
## Removing punctuation...
## Removing stopwords...
## Removing numbers...
## Stemming...
## Creating Output...

out <- prepDocuments(temp_first_ten$documents, temp_first_ten$vocab, temp_first_ten$meta)

## Removing 2110 of 4415 terms (2110 of 14305 tokens) due to frequency
## Your corpus now has 13 documents, 2305 terms and 12195 tokens.

# Run the model
# Should we change the value of K
model.stm <- stm(out$documents, out$vocab, K = 10, prevalence = ~ party,
                 data = out$meta, max.em.its = 10)

## Beginning Spectral Initialization
##    Calculating the gram matrix...
##    Finding anchor words...
##    ..........
##    Recovering initialization...
##    .....................
## Initialization complete.
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 1 (approx. per word bound = -6.617)
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 2 (approx. per word bound = -6.555, relative change = 9.262e-03)
## .............
## Completed E-Step (0 seconds).
```

```
## Completed M-Step.
## Completing Iteration 3 (approx. per word bound = -6.547, relative change = 1.321e-03)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 4 (approx. per word bound = -6.544, relative change = 4.609e-04)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 5 (approx. per word bound = -6.542, relative change = 2.994e-04)
## Topic 1: will, program, nation, peac, new
##  Topic 2: must, will, nation, can, everi
##  Topic 3: will, nation, american, can, new
##  Topic 4: world, will, must, nation, peac
##  Topic 5: nation, will, world, secur, congress
##  Topic 6: nation, program, new, year, feder
##  Topic 7: nation, year, will, world, tax
##  Topic 8: nation, program, year, will, must
##  Topic 9: nation, will, world, must, year
##  Topic 10: will, nation, new, year, world
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 6 (approx. per word bound = -6.540, relative change = 2.148e-04)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 7 (approx. per word bound = -6.539, relative change = 1.586e-04)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 8 (approx. per word bound = -6.539, relative change = 1.131e-04)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 9 (approx. per word bound = -6.538, relative change = 5.939e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Model Terminated Before Convergence Reached

model.stm <- stm(out$documents, out$vocab, K = 10, prevalence = ~ party,
                 data = out$meta)

## Beginning Spectral Initialization
##   Calculating the gram matrix...
##   Finding anchor words...
##    .........
##   Recovering initialization...
##    ....................
## Initialization complete.
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
```

```
## Completing Iteration 1 (approx. per word bound = -6.617)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 2 (approx. per word bound = -6.555, relative change = 9.262e-03)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 3 (approx. per word bound = -6.547, relative change = 1.321e-03)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 4 (approx. per word bound = -6.544, relative change = 4.609e-04)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 5 (approx. per word bound = -6.542, relative change = 2.994e-04)
## Topic 1: will, program, nation, peac, new
##  Topic 2: must, will, nation, can, everi
##  Topic 3: will, nation, american, can, new
##  Topic 4: world, will, must, nation, peac
##  Topic 5: nation, will, world, secur, congress
##  Topic 6: nation, program, new, year, feder
##  Topic 7: nation, year, will, world, tax
##  Topic 8: nation, program, year, will, must
##  Topic 9: nation, will, world, must, year
##  Topic 10: will, nation, new, year, world
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 6 (approx. per word bound = -6.540, relative change = 2.148e-04)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 7 (approx. per word bound = -6.539, relative change = 1.586e-04)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 8 (approx. per word bound = -6.539, relative change = 1.131e-04)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 9 (approx. per word bound = -6.538, relative change = 5.939e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 10 (approx. per word bound = -6.538, relative change = 4.784e-05)
## Topic 1: will, nation, program, peac, new
##  Topic 2: must, will, nation, can, everi
##  Topic 3: will, nation, american, can, new
##  Topic 4: world, will, peac, must, nation
##  Topic 5: nation, will, world, secur, congress
##  Topic 6: nation, program, new, year, feder
```

```
##  Topic 7: nation, year, will, world, tax
##  Topic 8: nation, program, year, will, must
##  Topic 9: nation, will, world, must, year
##  Topic 10: will, nation, new, year, world
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 11 (approx. per word bound = -6.538, relative change = 4.284e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 12 (approx. per word bound = -6.537, relative change = 4.255e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 13 (approx. per word bound = -6.537, relative change = 4.027e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 14 (approx. per word bound = -6.537, relative change = 3.777e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 15 (approx. per word bound = -6.537, relative change = 3.587e-05)
## Topic 1: will, nation, program, peac, new
##  Topic 2: must, will, nation, can, everi
##  Topic 3: will, nation, american, can, new
##  Topic 4: world, peac, will, must, nation
##  Topic 5: nation, will, world, secur, congress
##  Topic 6: nation, program, new, year, feder
##  Topic 7: nation, year, will, world, tax
##  Topic 8: nation, program, will, year, must
##  Topic 9: nation, will, world, must, year
##  Topic 10: will, nation, new, year, world
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 16 (approx. per word bound = -6.536, relative change = 3.375e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 17 (approx. per word bound = -6.536, relative change = 2.938e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 18 (approx. per word bound = -6.536, relative change = 2.606e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 19 (approx. per word bound = -6.536, relative change = 2.588e-05)
## ............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 20 (approx. per word bound = -6.536, relative change = 2.731e-05)
```

```
## Topic 1: will, nation, program, peac, new
##  Topic 2: must, will, nation, can, year
##  Topic 3: will, nation, american, can, new
##  Topic 4: world, peac, will, must, nation
##  Topic 5: nation, will, world, secur, congress
##  Topic 6: nation, program, new, year, feder
##  Topic 7: nation, year, will, world, tax
##  Topic 8: nation, will, program, year, must
##  Topic 9: nation, will, world, must, year
##  Topic 10: will, nation, new, year, world
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 21 (approx. per word bound = -6.536, relative change = 2.774e-05)
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 22 (approx. per word bound = -6.535, relative change = 2.164e-05)
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 23 (approx. per word bound = -6.535, relative change = 1.431e-05)
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 24 (approx. per word bound = -6.535, relative change = 1.170e-05)
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 25 (approx. per word bound = -6.535, relative change = 1.340e-05)
## Topic 1: will, nation, program, peac, new
##  Topic 2: must, will, nation, can, new
##  Topic 3: will, nation, american, can, new
##  Topic 4: world, peac, will, must, nation
##  Topic 5: nation, will, world, secur, congress
##  Topic 6: nation, program, new, year, feder
##  Topic 7: nation, year, will, world, tax
##  Topic 8: nation, will, program, year, must
##  Topic 9: nation, will, world, must, year
##  Topic 10: will, nation, new, year, world
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 26 (approx. per word bound = -6.535, relative change = 1.291e-05)
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 27 (approx. per word bound = -6.535, relative change = 1.373e-05)
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 28 (approx. per word bound = -6.535, relative change = 1.538e-05)
## .............
## Completed E-Step (0 seconds).
```

```
## Completed M-Step.
## Completing Iteration 29 (approx. per word bound = -6.535, relative change = 1.671e-05)
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 30 (approx. per word bound = -6.535, relative change = 1.845e-05)
## Topic 1: will, nation, program, peac, new
##  Topic 2: must, will, nation, can, new
##  Topic 3: will, nation, american, can, new
##  Topic 4: world, peac, must, will, militari
##  Topic 5: nation, will, world, secur, congress
##  Topic 6: nation, program, new, year, feder
##  Topic 7: nation, year, will, world, tax
##  Topic 8: nation, will, program, year, must
##  Topic 9: nation, will, world, must, year
##  Topic 10: will, new, nation, year, world
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 31 (approx. per word bound = -6.535, relative change = 1.500e-05)
## .............
## Completed E-Step (0 seconds).
## Completed M-Step.
## Model Converged

labelTopics(model.stm)

## Topic 1 Top Words:
##     Highest Prob: will, nation, program, peac, new, congress, year
##     FREX: accomplish, farmer, mention, subject, wartim, ever, thing
##     Lift: pictur, caught, modest, single-mind, wagehour, wartim, plain
##     Score: pictur, mention, wartim, accomplish, plain, refer, debt
## Topic 2 Top Words:
##     Highest Prob: must, will, nation, can, new, year, everi
##     FREX: lack, session, unemploy, ask, gold, let, job
##     Lift: hate, broader, jobless, lowest, nam, squalor, staff
##     Score: hate, lack, gold, tool, whatev, convent, bill
## Topic 3 Top Words:
##     Highest Prob: will, nation, american, can, new, year, world
##     FREX: presid, citi, man, centuri, learn, begin, seek
##     Lift: hill, elector, street, came, legaci, learn, recreat
##     Score: hill, learn, centuri, dream, recreat, river, shape
## Topic 4 Top Words:
##     Highest Prob: world, peac, must, will, militari, nation, defens
##     FREX: missil, soviet, militari, real, part, scienc, action
##     Lift: shield, concentr, jupit, solid, thor, featur, tempt
##     Score: shield, ballist, convict, extra, concentr, intellig, genuin
## Topic 5 Top Words:
##     Highest Prob: nation, will, world, secur, congress, peopl, govern
##     FREX: vital, principl, domest, agenc, money, affair, prosper
##     Lift: entiti, scene, assert, attitud, durabl, uphold, link
##     Score: entiti, assert, scene, stress, attitud, agenc, document
## Topic 6 Top Words:
##     Highest Prob: nation, program, new, year, feder, american, state
```
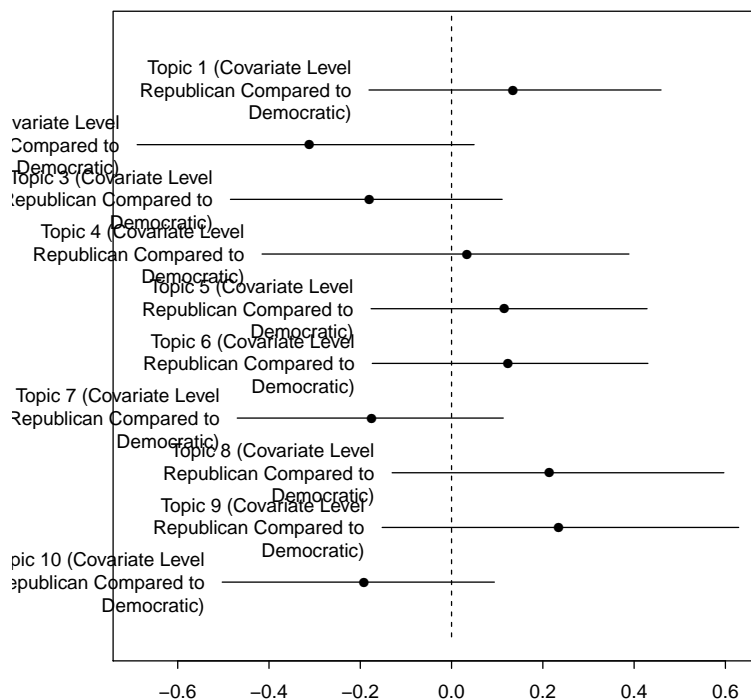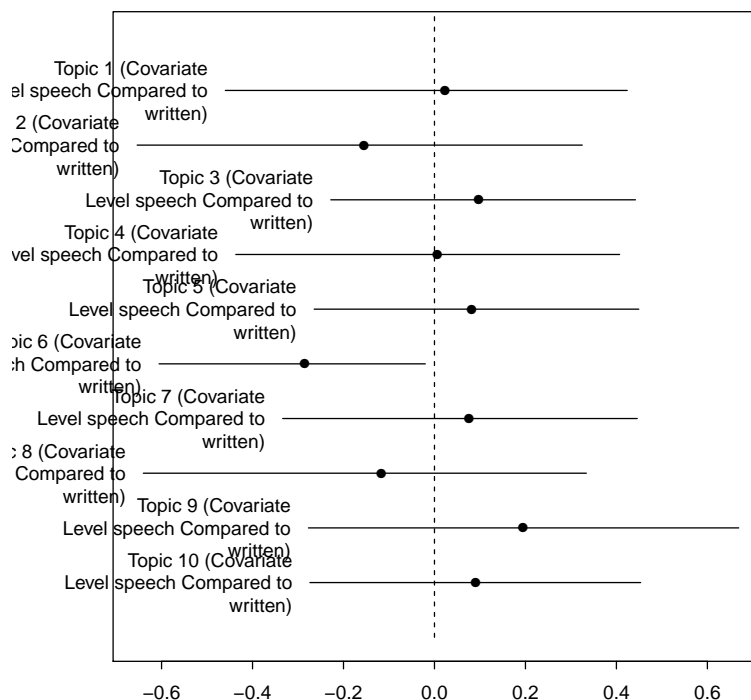
```
##      FREX: eight, veteran, sinc, percent, project, advanc, agenc
##      Lift: largest, eight, wheat, properti, elig, becam, veteran
##      Score: largest, eight, veteran, percent, wheat, agenc, becam
## Topic 7 Top Words:
##      Highest Prob: nation, year, will, world, tax, free, can
##      FREX: allianc, nuclear, end, west, tax, doubt, corpor
##      Lift: wind, viet-nam, door, victori, sudden, dead, dentist
##      Score: wind, choic, viet-nam, victori, cuba, worldwid, choos
## Topic 8 Top Words:
##      Highest Prob: nation, will, program, year, must, feder, govern
##      FREX: feder, adequ, legisl, forward, recommend, shall, measur
##      Lift: old-ag, contributori, harbor, self-govern, simpl, turkey, advisori
##      Score: old-ag, legisl, mainten, survivor, feder, district, vigor
## Topic 9 Top Words:
##      Highest Prob: nation, will, world, must, year, can, freedom
##      FREX: expenditur, price, determin, greater, cost, destruct, freedom
##      Lift: everyth, announc, gigant, interpol, outgo, twenti, unremit
##      Score: everyth, intellectu, ballist, court, treasuri, rule, solemn
## Topic 10 Top Words:
##      Highest Prob: will, new, nation, year, world, unit, free
##      FREX: bill, urg, market, atlant, job, across, offer
##      Lift: sometim, offer, commonsens, imperfect, adult, glori, mine
##      Score: sometim, bill, across, offer, choic, choos, fortun

# topics 1-10
model.stm.ee <- estimateEffect(1:10 ~ party, model.stm, meta = out$meta)
# right associated with cov.value1
plot(model.stm.ee, "party", method="difference", cov.value1="Republican", cov.value2="Democratic")
```

```
model.stm.ee <- estimateEffect(1:10 ~ sotu_type, model.stm, meta = out$meta)
plot(model.stm.ee, "sotu_type", method="difference", cov.value1="speech", cov.value2="written")
```



```
########################################## last ten year of Vietnam War ##########################
# The following code relates to the last ten year of the Vietnam War
war_last_ten <- metadata[which((metadata$year>=1966) & (metadata$year<=1975)),c(colnames(metadata))]
# Preprocessing of the  data and make it corpus
corpus_sotu_last_ten <- corpus(war_last_ten, text_field = "text")
corpus_sotu_last_ten

## Corpus consisting of 11 documents and 4 docvars.
## text1 :
## " [ Delivered in person before a joint session at 9:04 p.m. ]..."
##
## text2 :
## " [Delivered in person before a joint session at 9:33 p.m.]  ..."
##
## text3 :
## " [Delivered in person before a joint session at 9:05 p.m.]  ..."
##
## text4 :
## "Mr. Speaker, Mr. President, Members of the Congress and my f..."
##
## text5 :
## " Mr. Speaker, Mr. President, my colleagues in the Congress, ..."
##
## text6 :
## " Mr. Speaker, Mr. President, my colleagues in the Congress, ..."
##
## [ reached max_ndoc ... 5 more documents ]
```

```
#Some common pre-processing, remove the punctuation and numbers
toks_last_ten <- tokens(corpus_sotu_last_ten, remove_punct = TRUE, remove_numbers=TRUE)
toks_last_ten <- tokens_wordstem(toks_last_ten)
toks_last_ten <- tokens_select(toks_last_ten,  stopwords("en"), selection = "remove")
dfm <- dfm(toks_last_ten)
dfm

## Document-feature matrix of: 11 documents, 4,086 features (71.76% sparse) and 4 docvars.
##        features
## docs    deliv person befor joint session p.m mr speaker presid member
##    text1    1      2     8     1       1   1  2       1      9      6
##    text2    1      3     9     1       2   1  3       1     14      5
##    text3    1      3     8     1       3   1  3       1      6      1
##    text4    0      5     4     0       0   0  4       3     15      5
##    text5    0      2     3     1       2   0  2       1      4      2
##    text6    0      2     4     0       2   0  4       5      3      4
## [ reached max_ndoc ... 5 more documents, reached max_nfeat ... 4,076 more features ]

#Create a document feature matrix (dfm) and trim it with words appeared at least 5%
toks_last_ten <- corpus_sotu %>%
  tokens()
dfm_last_ten <- dfm(toks_last_ten)
dfm_trimmed_last_ten <- dfm_trim(dfm_last_ten, min_docfreq = 0.05, docfreq_type = "prop")
dfm_trimmed_last_ten

## Document-feature matrix of: 24 documents, 5,099 features (72.87% sparse) and 4 docvars.
##        features
## docs    mr    . president   , speaker members  of the congress  :
##    text1  2 333         3 368       1      2 370 440       24 11
##    text2  0  52         0  44       0      0  43  79        6  2
##    text3  0 381         0 336       0      0 420 609       29 10
##    text4  0 190         0 178       0      1 207 278       18  6
##    text5  2 242         2 260       1      3 231 312       10 22
##    text6  2 260         3 198       1      2 239 291       16  8
## [ reached max_ndoc ... 18 more documents, reached max_nfeat ... 5,089 more features ]

# word cloud for the first ten years of the Vietnam War
textplot_wordcloud(dfm_trimmed_last_ten, col="purple")
```

```r
# Find words that are distinctive between newsletters written by Democrats and Republicans
terms_last_ten_s <- clusterFightinWords(dfm_trimmed_last_ten,war_last_ten$sotu_type == "speech")
sort(terms_last_ten_s, decreasing=T)[1:10] #nation,people
```

```
##          of      peace     federal   necessary       1974      energy        made
##    1.759510    1.694130    1.669265    1.563778    1.560714    1.536161    1.501099
##         oil legislation   reduction
##    1.465342    1.314701    1.299051
```

```r
terms_last_ten_s <- clusterFightinWords(dfm_trimmed_last_ten,war_first_ten$sotu_type == "written")
sort(terms_last_ten_s, decreasing=T)[1:10]#federal, administration, veterans, education
```

```
##        been     program        1953         has        also     sharing     welfare     revenue   influence
##    5.177928    4.578795    3.740694    3.085252    2.930361    2.777721    2.718907    2.714781    2.594274
##        last
##    2.518789
```

```r
############# LDA ############

#Run LDA using quanteda
lda <- textmodel_lda(dfm_trimmed_last_ten, k = 10)
#Most likely term for each topic
lda <- terms(lda, 10)
lda.terms
```

```
##      topic1       topic2   topic3    topic4       topic5      topic6    topic7
## [1,] "must"       "-"      "-"       "energy"     "free"      "-"       "the"
## [2,] "government" "our"    "$"       "federal"    "world"     "also"    "."
## [3,] "justice"    "or"     "billion" "$"          "defense"   "which"   ","
## [4,] "citizens"   "and"    "percent" "would"      "nations"   "on"      "of"
## [5,] "recommend"  "but"    "million" "will"       "peace"     "new"     "and"
```

```
##  [6,] "shall"          "must"   "last"   "-"            "security" "for"     "to"
##  [7,] "human"          "not"    "very"   "oil"          "military" "have"    "in"
##  [8,] "laws"           "its"    "act"    "legislation" "all"       "reform"  "a"
##  [9,] ";"              "his"    "who"    "economic"     "economic" "my"      "our"
## [10,] "peoples"        "every" "home"    "needs"        "freedom"  "percent" "we"
##       topic8            topic9          topic10
##  [1,] "been"            "one"           "we"
##  [2,] "have"            "position"      "that"
##  [3,] "has"             "scientific"    "i"
##  [4,] "government"      "missiles"      "will"
##  [5,] "federal"         "if"            "have"
##  [6,] "administration" "development"   "america"
##  [7,] "program"         "education"     "great"
##  [8,] "also"            "greater"       "government"
##  [9,] "system"          "missile"       "you"
## [10,] "development"     "technology"    "us"
```

```r
#Topical content matrix
mu_last_ten <- lda$phi
```

```
## Error in lda$phi:  $ operator is invalid for atomic vectors
```

```r
dim(mu_last_ten) #10 topics, 5099 words
```

```
## Error in eval(expr, envir, enclos):  object 'mu_last_ten' not found
```

```r
mu_last_ten[1:10,1:20]
```

```
## Error in eval(expr, envir, enclos):  object 'mu_last_ten' not found
```

```r
#Most representative words in Topic 1-10
for (i in 1:10){
  print(mu_last_ten[i,][order(mu_last_ten[i,], decreasing=T)][1:10])
  print("####################################################")
}
```

```
## Error in print(mu_last_ten[i, ][order(mu_last_ten[i, ], decreasing = T)][1:10]):  object
'mu_last_ten' not found
```

```r
#Topical prevalence matrix
pi <- lda_last_ten$theta
```

```
## Error in lda_last_ten$theta:  $ operator is invalid for atomic vectors
```

```r
dim(pi) #24 10
```

```
## [1] 24 10
```

```r
#Most representative documents in Topic 1-10
for (i in 1:10){
  print(war_last_ten[order(pi[i,],decreasing=T),])
  print("####################################################")
}
```

```
## # A tibble: 10 x 5
##    president            year party      sotu_type text
##    <chr>               <dbl> <chr>      <chr>      <chr>
##  1 Richard M. Nixon     1972 Republican speech     "\n\nMr. Speaker, Mr. President, my colle~
```

```
##  2 Richard M. Nixon   1972 Republican written   "\n\nTo the Congress of the United States~
##  3 Lyndon B. Johnson  1966 Democratic speech    "\n\n[ Delivered in person before a joint~
##  4 Richard M. Nixon   1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  5 Lyndon B. Johnson  1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  6 Lyndon B. Johnson  1968 Democratic speech    "\n\n[Delivered in person before a joint ~
##  7 Richard M. Nixon   1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  8 Richard M. Nixon   1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
##  9 Richard M. Nixon   1974 Republican written   "To the Congress of the United States: \n~
## 10 Lyndon B. Johnson  1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president          year party      sotu_type text
##    <chr>             <dbl> <chr>      <chr>     <chr>
##  1 Richard M. Nixon   1972 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  2 Richard M. Nixon   1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  3 Richard M. Nixon   1972 Republican written   "\n\nTo the Congress of the United States~
##  4 Richard M. Nixon   1974 Republican written   "To the Congress of the United States: \n~
##  5 Lyndon B. Johnson  1966 Democratic speech    "\n\n[ Delivered in person before a joint~
##  6 Lyndon B. Johnson  1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  7 Lyndon B. Johnson  1968 Democratic speech    "\n\n[Delivered in person before a joint ~
##  8 Richard M. Nixon   1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
##  9 Lyndon B. Johnson  1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
## 10 Richard M. Nixon   1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president          year party      sotu_type text
##    <chr>             <dbl> <chr>      <chr>     <chr>
##  1 Richard M. Nixon   1972 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  2 Richard M. Nixon   1972 Republican written   "\n\nTo the Congress of the United States~
##  3 Lyndon B. Johnson  1966 Democratic speech    "\n\n[ Delivered in person before a joint~
##  4 Richard M. Nixon   1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  5 Lyndon B. Johnson  1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  6 Richard M. Nixon   1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  7 Richard M. Nixon   1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
##  8 Lyndon B. Johnson  1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
##  9 Richard M. Nixon   1974 Republican written   "To the Congress of the United States: \n~
## 10 Lyndon B. Johnson  1968 Democratic speech    "\n\n[Delivered in person before a joint ~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president          year party      sotu_type text
##    <chr>             <dbl> <chr>      <chr>     <chr>
##  1 Richard M. Nixon   1972 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  2 Richard M. Nixon   1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  3 Lyndon B. Johnson  1966 Democratic speech    "\n\n[ Delivered in person before a joint~
##  4 Richard M. Nixon   1972 Republican written   "\n\nTo the Congress of the United States~
##  5 Lyndon B. Johnson  1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  6 Lyndon B. Johnson  1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
##  7 Richard M. Nixon   1974 Republican written   "To the Congress of the United States: \n~
##  8 Richard M. Nixon   1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  9 Lyndon B. Johnson  1968 Democratic speech    "\n\n[Delivered in person before a joint ~
## 10 Richard M. Nixon   1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
## [1] "#####################################################"
## # A tibble: 10 x 5
```

```
##    president         year party      sotu_type text
##    <chr>            <dbl> <chr>      <chr>     <chr>
##  1 Richard M. Nixon  1972 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  2 Richard M. Nixon  1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  3 Richard M. Nixon  1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
##  4 Richard M. Nixon  1974 Republican written   "To the Congress of the United States: \n~
##  5 Lyndon B. Johnson 1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  6 Lyndon B. Johnson 1966 Democratic speech    "\n\n[ Delivered in person before a joint~
##  7 Richard M. Nixon  1972 Republican written   "\n\nTo the Congress of the United States~
##  8 Richard M. Nixon  1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  9 Lyndon B. Johnson 1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
## 10 Lyndon B. Johnson 1968 Democratic speech    "\n\n[Delivered in person before a joint ~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president         year party      sotu_type text
##    <chr>            <dbl> <chr>      <chr>     <chr>
##  1 Richard M. Nixon  1972 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  2 Richard M. Nixon  1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  3 Richard M. Nixon  1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
##  4 Lyndon B. Johnson 1966 Democratic speech    "\n\n[ Delivered in person before a joint~
##  5 Richard M. Nixon  1972 Republican written   "\n\nTo the Congress of the United States~
##  6 Lyndon B. Johnson 1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  7 Richard M. Nixon  1974 Republican written   "To the Congress of the United States: \n~
##  8 Lyndon B. Johnson 1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
##  9 Richard M. Nixon  1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
## 10 Lyndon B. Johnson 1968 Democratic speech    "\n\n[Delivered in person before a joint ~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president         year party      sotu_type text
##    <chr>            <dbl> <chr>      <chr>     <chr>
##  1 Richard M. Nixon  1972 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  2 Richard M. Nixon  1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
##  3 Richard M. Nixon  1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  4 Lyndon B. Johnson 1966 Democratic speech    "\n\n[ Delivered in person before a joint~
##  5 Lyndon B. Johnson 1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  6 Richard M. Nixon  1974 Republican written   "To the Congress of the United States: \n~
##  7 Richard M. Nixon  1972 Republican written   "\n\nTo the Congress of the United States~
##  8 Lyndon B. Johnson 1968 Democratic speech    "\n\n[Delivered in person before a joint ~
##  9 Lyndon B. Johnson 1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
## 10 Richard M. Nixon  1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
## [1] "#####################################################"
## # A tibble: 10 x 5
##    president         year party      sotu_type text
##    <chr>            <dbl> <chr>      <chr>     <chr>
##  1 Richard M. Nixon  1972 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  2 Lyndon B. Johnson 1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  3 Richard M. Nixon  1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
##  4 Richard M. Nixon  1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  5 Richard M. Nixon  1972 Republican written   "\n\nTo the Congress of the United States~
##  6 Richard M. Nixon  1974 Republican written   "To the Congress of the United States: \n~
##  7 Lyndon B. Johnson 1968 Democratic speech    "\n\n[Delivered in person before a joint ~
##  8 Lyndon B. Johnson 1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
##  9 Richard M. Nixon  1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
## 10 Lyndon B. Johnson 1966 Democratic speech    "\n\n[ Delivered in person before a joint~
```

```
## [1] "##################################################"
## # A tibble: 10 x 5
##    president          year party      sotu_type text
##    <chr>             <dbl> <chr>      <chr>     <chr>
##  1 Richard M. Nixon   1972 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  2 Richard M. Nixon   1972 Republican written   "\n\nTo the Congress of the United States~
##  3 Richard M. Nixon   1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  4 Richard M. Nixon   1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
##  5 Lyndon B. Johnson  1968 Democratic speech    "\n\n[Delivered in person before a joint ~
##  6 Lyndon B. Johnson  1966 Democratic speech    "\n\n[ Delivered in person before a joint~
##  7 Richard M. Nixon   1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  8 Lyndon B. Johnson  1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  9 Lyndon B. Johnson  1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
## 10 Richard M. Nixon   1974 Republican written   "To the Congress of the United States: \n~
## [1] "##################################################"
## # A tibble: 10 x 5
##    president          year party      sotu_type text
##    <chr>             <dbl> <chr>      <chr>     <chr>
##  1 Richard M. Nixon   1972 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  2 Lyndon B. Johnson  1967 Democratic speech    "\n\n[Delivered in person before a joint ~
##  3 Lyndon B. Johnson  1968 Democratic speech    "\n\n[Delivered in person before a joint ~
##  4 Richard M. Nixon   1970 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  5 Richard M. Nixon   1974 Republican written   "To the Congress of the United States: \n~
##  6 Richard M. Nixon   1974 Republican speech    "Mr. Speaker, Mr. President, my colleague~
##  7 Richard M. Nixon   1972 Republican written   "\n\nTo the Congress of the United States~
##  8 Richard M. Nixon   1971 Republican speech    "\n\nMr. Speaker, Mr. President, my colle~
##  9 Lyndon B. Johnson  1969 Democratic written   "Mr. Speaker, Mr. President, Members of t~
## 10 Lyndon B. Johnson  1966 Democratic speech    "\n\n[ Delivered in person before a joint~
## [1] "##################################################"


############ STM ###########
# Structural Topic Model
#STM
#library(tm)
temp_last_ten <-textProcessor(documents=war_last_ten$text,metadata=war_last_ten)

## Building corpus...
## Converting to Lower Case...
## Removing punctuation...
## Removing stopwords...
## Removing numbers...
## Stemming...
## Creating Output...

out <- prepDocuments(temp_last_ten$documents, temp_last_ten$vocab, temp_last_ten$meta)

## Removing 2244 of 4434 terms (2244 of 12768 tokens) due to frequency
## Your corpus now has 11 documents, 2190 terms and 10524 tokens.

# Run the model
# Should we change the value of K
model.stm <- stm(out$documents, out$vocab, K = 10, prevalence = ~ party,
                 data = out$meta, max.em.its = 10)

## Beginning Spectral Initialization
```

```
##   Calculating the gram matrix...
##   Finding anchor words...
##   ..........
##   Recovering initialization...
##   ....................
## Initialization complete.
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 1 (approx. per word bound = -6.503)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 2 (approx. per word bound = -6.465, relative change = 5.794e-03)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 3 (approx. per word bound = -6.454, relative change = 1.759e-03)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 4 (approx. per word bound = -6.450, relative change = 6.548e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 5 (approx. per word bound = -6.448, relative change = 3.245e-04)
## Topic 1: govern, will, peopl, can, congress
##  Topic 2: will, tax, year, energi, world
##  Topic 3: year, congress, will, nation, think
##  Topic 4: year, will, american, america, peac
##  Topic 5: will, peopl, year, nation, vietnam
##  Topic 6: new, will, america, nation, year
##  Topic 7: will, year, can, congress, billion
##  Topic 8: will, year, new, nation, congress
##  Topic 9: feder, will, program, also, new
##  Topic 10: will, year, nation, can, must
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 6 (approx. per word bound = -6.446, relative change = 1.965e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 7 (approx. per word bound = -6.445, relative change = 1.527e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 8 (approx. per word bound = -6.444, relative change = 1.494e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 9 (approx. per word bound = -6.443, relative change = 1.493e-04)
## ..........
## Completed E-Step (0 seconds).
```

```
## Completed M-Step.
## Model Terminated Before Convergence Reached

model.stm <- stm(out$documents, out$vocab, K = 10, prevalence = ~ party,
                 data = out$meta)

## Beginning Spectral Initialization
##    Calculating the gram matrix...
##    Finding anchor words...
##    ..........
##    Recovering initialization...
##    ...................
## Initialization complete.
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 1 (approx. per word bound = -6.503)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 2 (approx. per word bound = -6.465, relative change = 5.794e-03)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 3 (approx. per word bound = -6.454, relative change = 1.759e-03)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 4 (approx. per word bound = -6.450, relative change = 6.548e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 5 (approx. per word bound = -6.448, relative change = 3.245e-04)
## Topic 1: govern, will, peopl, can, congress
##  Topic 2: will, tax, year, energi, world
##  Topic 3: year, congress, will, nation, think
##  Topic 4: year, will, american, america, peac
##  Topic 5: will, peopl, year, nation, vietnam
##  Topic 6: new, will, america, nation, year
##  Topic 7: will, year, can, congress, billion
##  Topic 8: will, year, new, nation, congress
##  Topic 9: feder, will, program, also, new
##  Topic 10: will, year, nation, can, must
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 6 (approx. per word bound = -6.446, relative change = 1.965e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 7 (approx. per word bound = -6.445, relative change = 1.527e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
```

```
## Completing Iteration 8 (approx. per word bound = -6.444, relative change = 1.494e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 9 (approx. per word bound = -6.443, relative change = 1.493e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 10 (approx. per word bound = -6.442, relative change = 1.453e-04)
## Topic 1: govern, will, peopl, can, congress
##  Topic 2: will, tax, year, energi, new
##  Topic 3: year, congress, will, think, nation
##  Topic 4: year, will, america, american, peac
##  Topic 5: will, peopl, year, nation, vietnam
##  Topic 6: new, will, america, nation, year
##  Topic 7: will, year, can, congress, billion
##  Topic 8: will, year, new, nation, can
##  Topic 9: will, feder, program, year, also
##  Topic 10: will, nation, year, can, must
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 11 (approx. per word bound = -6.442, relative change = 1.336e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 12 (approx. per word bound = -6.441, relative change = 1.276e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 13 (approx. per word bound = -6.440, relative change = 1.200e-04)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 14 (approx. per word bound = -6.439, relative change = 9.608e-05)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 15 (approx. per word bound = -6.439, relative change = 6.829e-05)
## Topic 1: govern, will, peopl, can, congress
##  Topic 2: will, tax, year, energi, new
##  Topic 3: year, congress, will, think, nation
##  Topic 4: year, will, america, american, peac
##  Topic 5: will, peopl, year, nation, vietnam
##  Topic 6: new, will, america, nation, year
##  Topic 7: will, year, can, congress, billion
##  Topic 8: will, year, new, can, nation
##  Topic 9: will, feder, year, program, increas
##  Topic 10: will, nation, year, can, must
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 16 (approx. per word bound = -6.439, relative change = 5.036e-05)
```

```
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 17 (approx. per word bound = -6.438, relative change = 3.980e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 18 (approx. per word bound = -6.438, relative change = 3.377e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 19 (approx. per word bound = -6.438, relative change = 2.836e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 20 (approx. per word bound = -6.438, relative change = 2.457e-05)
## Topic 1: govern, will, peopl, can, congress
##  Topic 2: will, tax, year, energi, new
##  Topic 3: year, congress, will, think, nation
##  Topic 4: year, will, america, american, peac
##  Topic 5: will, peopl, year, nation, vietnam
##  Topic 6: new, will, america, nation, year
##  Topic 7: will, year, can, congress, billion
##  Topic 8: will, year, new, can, nation
##  Topic 9: will, feder, year, program, increas
##  Topic 10: will, nation, year, can, must
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 21 (approx. per word bound = -6.438, relative change = 2.214e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 22 (approx. per word bound = -6.437, relative change = 2.063e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 23 (approx. per word bound = -6.437, relative change = 1.821e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 24 (approx. per word bound = -6.437, relative change = 1.602e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 25 (approx. per word bound = -6.437, relative change = 1.378e-05)
## Topic 1: govern, will, peopl, can, congress
##  Topic 2: will, tax, year, energi, new
##  Topic 3: year, congress, will, think, nation
##  Topic 4: year, will, america, american, peac
##  Topic 5: will, peopl, year, nation, vietnam
##  Topic 6: new, will, america, nation, year
##  Topic 7: year, will, can, congress, billion
##  Topic 8: will, year, new, can, nation
```

```
##  Topic 9: will, feder, year, program, increas
##  Topic 10: will, nation, year, can, must
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 26 (approx. per word bound = -6.437, relative change = 1.254e-05)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 27 (approx. per word bound = -6.437, relative change = 1.139e-05)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 28 (approx. per word bound = -6.437, relative change = 1.139e-05)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 29 (approx. per word bound = -6.437, relative change = 1.118e-05)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 30 (approx. per word bound = -6.437, relative change = 1.138e-05)
## Topic 1: govern, will, peopl, can, congress
##  Topic 2: will, tax, year, energi, new
##  Topic 3: year, congress, will, think, nation
##  Topic 4: year, will, america, american, peac
##  Topic 5: will, peopl, year, nation, vietnam
##  Topic 6: new, will, america, nation, year
##  Topic 7: year, will, can, congress, billion
##  Topic 8: will, new, year, can, nation
##  Topic 9: will, feder, year, program, increas
##  Topic 10: will, nation, year, can, must
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 31 (approx. per word bound = -6.437, relative change = 1.317e-05)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 32 (approx. per word bound = -6.437, relative change = 1.358e-05)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 33 (approx. per word bound = -6.437, relative change = 1.378e-05)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 34 (approx. per word bound = -6.436, relative change = 1.403e-05)
## ..........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 35 (approx. per word bound = -6.436, relative change = 1.406e-05)
## Topic 1: govern, will, peopl, can, congress
##  Topic 2: will, tax, year, energi, new
```

```
##  Topic 3: year, congress, will, think, nation
##  Topic 4: year, will, america, american, peac
##  Topic 5: will, peopl, year, nation, vietnam
##  Topic 6: new, will, america, nation, year
##  Topic 7: year, will, can, congress, billion
##  Topic 8: will, new, year, can, program
##  Topic 9: will, feder, year, program, increas
##  Topic 10: will, nation, year, can, must
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 36 (approx. per word bound = -6.436, relative change = 1.383e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 37 (approx. per word bound = -6.436, relative change = 1.310e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 38 (approx. per word bound = -6.436, relative change = 1.417e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 39 (approx. per word bound = -6.436, relative change = 1.265e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 40 (approx. per word bound = -6.436, relative change = 1.218e-05)
## Topic 1: govern, will, peopl, can, congress
##  Topic 2: will, tax, year, energi, new
##  Topic 3: year, congress, will, think, nation
##  Topic 4: year, will, america, american, peac
##  Topic 5: will, peopl, year, nation, vietnam
##  Topic 6: new, will, america, nation, year
##  Topic 7: year, will, can, congress, billion
##  Topic 8: will, new, year, can, program
##  Topic 9: will, feder, year, program, increas
##  Topic 10: will, nation, year, can, must
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 41 (approx. per word bound = -6.436, relative change = 1.119e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Completing Iteration 42 (approx. per word bound = -6.436, relative change = 1.084e-05)
## ...........
## Completed E-Step (0 seconds).
## Completed M-Step.
## Model Converged

labelTopics(model.stm)

## Topic 1 Top Words:
```
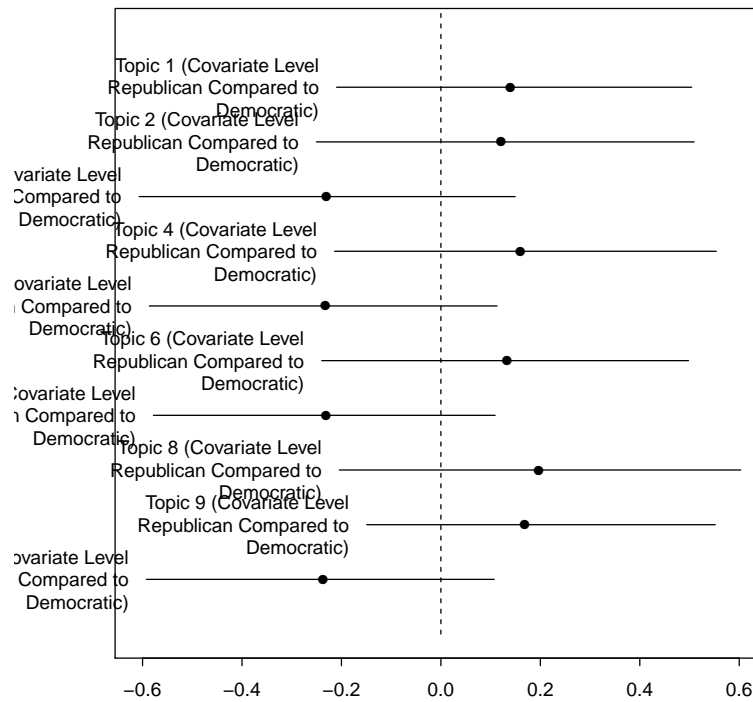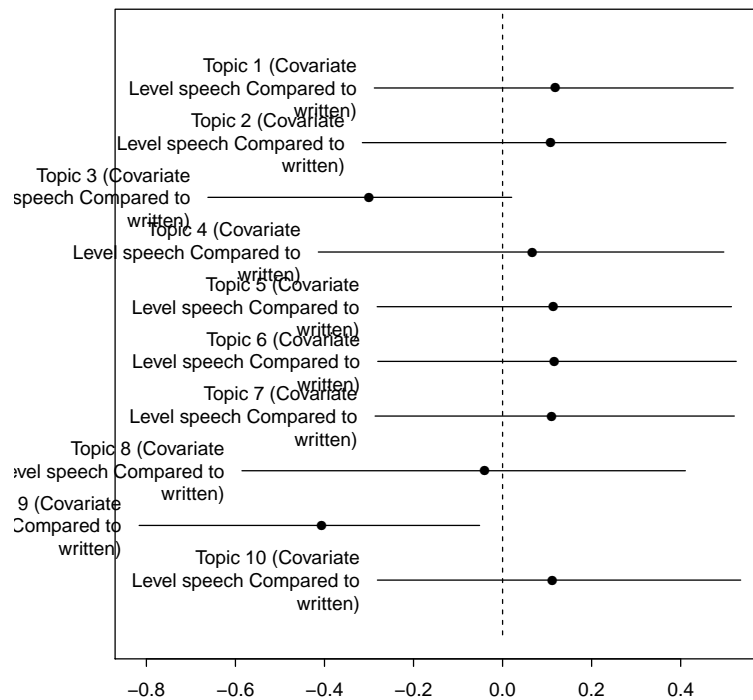
```
##     Highest Prob: govern, will, peopl, can, congress, nation, new
##     FREX: door, open, function, govern, depart, chanc, truli
##     Lift: function, door, rescu, rob, congratul, escap, russel
##     Score: function, vision, truli, depart, door, open, peacetim
## Topic 2 Top Words:
##     Highest Prob: will, tax, year, energi, new, world, program
##     FREX: oil, suppli, per, tax, energi, coal, depend
##     Lift: petroleum, attain, fee, happi, freshman, raw, synthet
##     Score: petroleum, oil, happi, coal, attain, suppli, freshman
## Topic 3 Top Words:
##     Highest Prob: year, congress, will, think, nation, american, believ
##     FREX: think, believ, commit, throughout, poverti, want, alreadi
##     Lift: nixon, cycl, glad, languag, commonsens, czechoslovakia, unanim
##     Score: nixon, tonight, poverti, salari, commonsens, czechoslovakia, unanim
## Topic 4 Top Words:
##     Highest Prob: year, will, america, american, peac, can, nation
##     FREX: toward, investig, five, energi, peac, messag, afford
##     Lift: investig, depress, recess, outrag, judiciari, freshmen, longest
##     Score: investig, privaci, recess, truli, tonight, strongest, messag
## Topic 5 Top Words:
##     Highest Prob: will, peopl, year, nation, vietnam, tonight, must
##     FREX: south, vietnam, principl, modern, north, desir, attack
##     Lift: servant, conquest, river, pleasur, soldier, ravag, path
##     Score: servant, tonight, north, aggress, fire, conquest, endur
## Topic 6 Top Words:
##     Highest Prob: new, will, america, nation, year, peopl, american
##     FREX: clean, seventi, water, sens, growth, decad, primari
##     Lift: spiritu, debt, suburb, blame, abolish, seventi, none
##     Score: spiritu, debt, blame, seventi, none, vision, sixti
## Topic 7 Top Words:
##     Highest Prob: year, will, can, congress, billion, nation, american
##     FREX: travel, bill, billion, pass, consum, expenditur, abund
##     Lift: ship, gold, ambassador, repeal, middle-incom, attorney, mandatori
##     Score: ship, polic, tonight, gold, travel, repeal, middle-incom
## Topic 8 Top Words:
##     Highest Prob: will, new, year, can, program, also, nation
##     FREX: influenc, research, school, technolog, expand, older, employ
##     Lift: alliesw, bind, blind, canada, commitmentsw, defense-rel, deterior
##     Score: influenc, blind, research, school, expand, reform, technolog
## Topic 9 Top Words:
##     Highest Prob: will, feder, year, program, new, increas, also
##     FREX: indian, environment, women, intern, veteran, assist, educ
##     Lift: alter, discourag, especi, fha, harvest, passeng, pleas
##     Score: mortgag, environment, oil, indian, veteran, ensur, reform
## Topic 10 Top Words:
##     Highest Prob: will, nation, year, can, must, american, congress
##     FREX: tri, succeed, whether, seem, month, control, sometim
##     Lift: sometim, asian, movement, compel, tactic, outlaw, surcharg
##     Score: sometim, tonight, asian, wish, polic, movement, poverti

# topics 1-10
model.stm.ee <- estimateEffect(1:10 ~ party, model.stm, meta = out$meta)
# right associated with cov.value1
plot(model.stm.ee, "party", method="difference", cov.value1="Republican", cov.value2="Democratic")
```

```r
model.stm.ee <- estimateEffect(1:10 ~ sotu_type, model.stm, meta = out$meta)
plot(model.stm.ee, "sotu_type", method="difference", cov.value1="speech", cov.value2="written")
```


The R session information (including the OS info, R version and all packages used):

```
sessionInfo()

## R version 3.6.2 (2019-12-12)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS  10.16
##
## Matrix products: default
## LAPACK: /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] seededlda_0.6.0         stm_1.3.6               quanteda.textplots_0.94
##  [4] quanteda_3.0.0          tokenizers_0.2.1        forcats_0.5.1
##  [7] stringr_1.4.0           dplyr_1.0.7             purrr_0.3.4
## [10] readr_2.1.2             tidyr_1.2.0             tibble_3.1.6
## [13] ggplot2_3.3.5           tidyverse_1.3.1
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.8          lubridate_1.8.0    lattice_0.20-38    digest_0.6.27
##  [5] assertthat_0.2.1    utf8_1.2.2         slam_0.1-50        R6_2.5.1
##  [9] cellranger_1.1.0    backports_1.4.1    reprex_2.0.1       evaluate_0.14
## [13] highr_0.8           httr_1.4.2         pillar_1.7.0       rlang_1.0.1
## [17] readxl_1.3.1        rstudioapi_0.13    data.table_1.14.2  Matrix_1.2-18
## [21] rmarkdown_2.7       bit_4.0.4          munsell_0.5.0      broom_0.7.12
## [25] xfun_0.22           compiler_3.6.2     modelr_0.1.8       pkgconfig_2.0.3
## [29] htmltools_0.5.1.1   tidyselect_1.1.1   matrixStats_0.61.0 fansi_1.0.2
## [33] crayon_1.4.2        tzdb_0.2.0         dbplyr_2.1.1       withr_2.4.3
## [37] SnowballC_0.7.0     grid_3.6.2         jsonlite_1.7.2     gtable_0.3.0
## [41] lifecycle_1.0.1     DBI_1.1.2          magrittr_2.0.1     scales_1.1.1
## [45] RcppParallel_5.1.5  cli_3.1.1          stringi_1.5.3      vroom_1.5.7
## [49] fs_1.5.2            NLP_0.2-1          xml2_1.3.3         ellipsis_0.3.2
## [53] stopwords_2.3       generics_0.1.2     vctrs_0.3.8        fastmatch_1.1-3
## [57] tools_3.6.2         bit64_4.0.5        glue_1.4.2         hms_1.1.1
## [61] yaml_2.2.1          parallel_3.6.2     tm_0.7-8           colorspace_2.0-2
## [65] rvest_1.0.2         knitr_1.31         haven_2.4.3

Sys.time()

## [1] "2022-05-24 02:44:55 PDT"
```